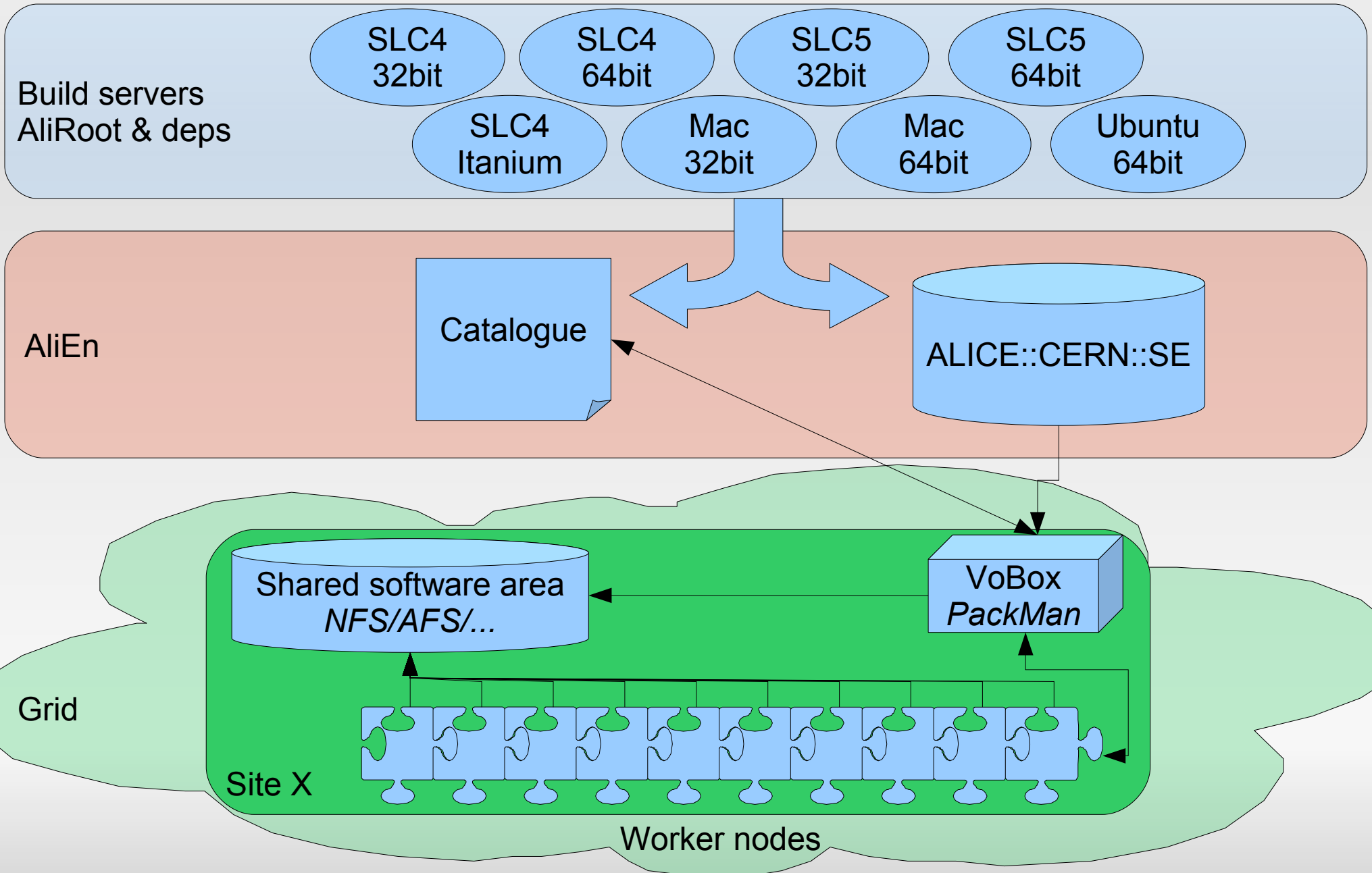


Torrent-based software distribution

Costin Grigoras
Pablo Saiz

Current way of distributing sw



Current way of distributing sw

Advantages

- A single service/site manages the installation of required packages

Disadvantages

- Shared software area is a single point of failure / bottleneck
- Difficult to update packages keeping the version number
- Need to keep a short list of active software packages

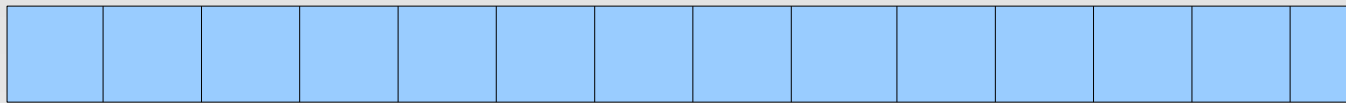
How can we avoid using a shared software area ?

- Worker nodes are independent
 - Self-consistent software packages are required
- No site-local software repository
 - Avoid overloading central software repositories
- Would be nice to be able to quickly update software packages if needed

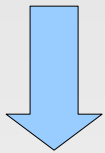
- We are trying to use BitTorrent technology to solve all the above

Preparing for torrent

package.tar.bz2



Chunks of equal size



package.tar.bz2.torrent

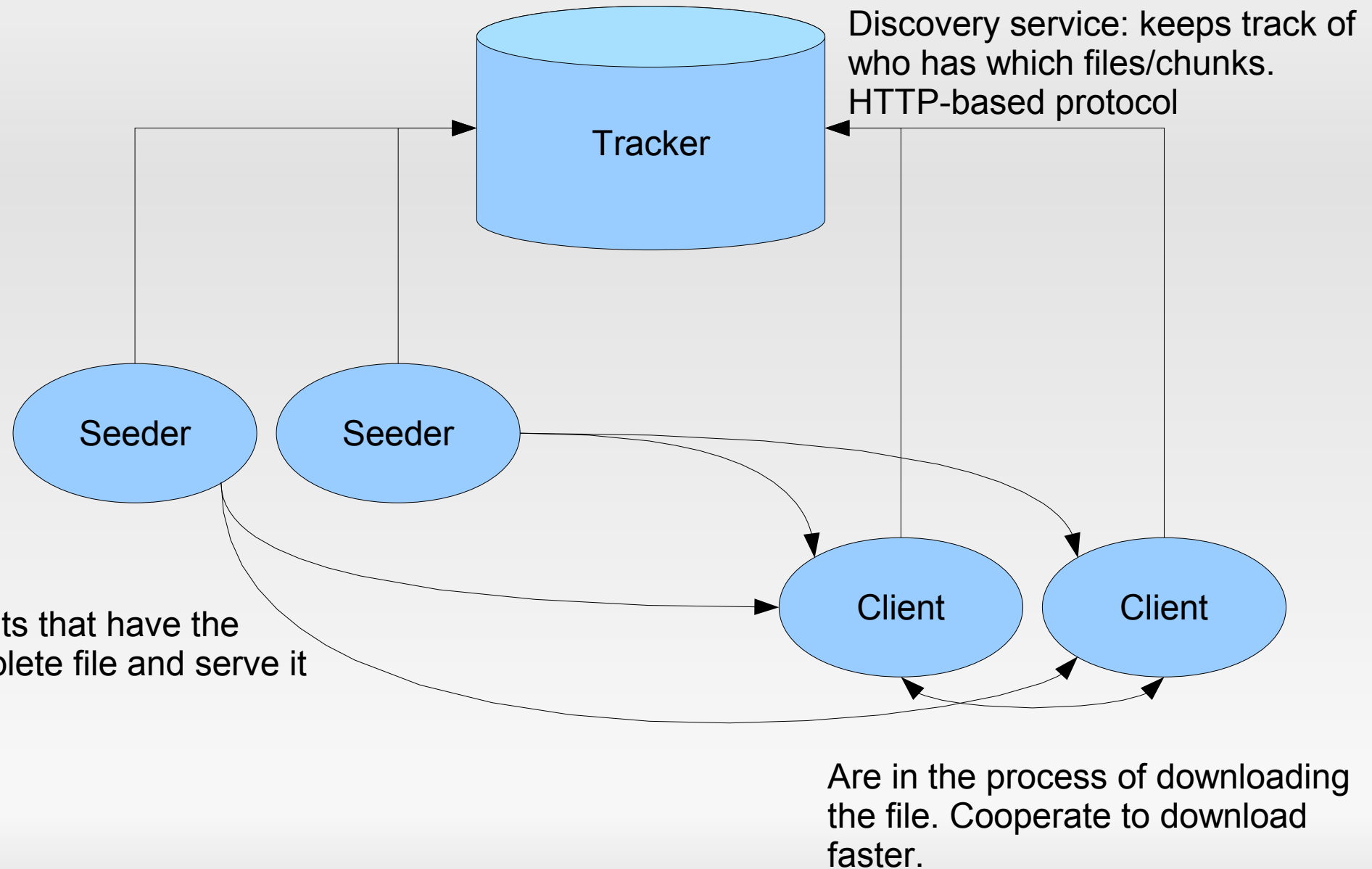


(tens of KB)

Metadata info of the original file:

- SHA1 hashes of chunks
- SHA1 hash of the entire file
 - * uniquely identifies the file
- Tracker location (entry point)

Data flow in torrent networks



Implementation in AliEn

Build servers
AliRoot & deps

SLC4
32bit

SLC4
64bit

SLC5
32bit

SLC5
64bit

SLC4
Itanium

Mac
32bit

Mac
64bit

Ubuntu
64bit

AliEn

Catalogue
torrent://...

<http://alitorrent.cern.ch>

Seeder
alitorrent:8092

Tracker
alitorrent:8088

VoBox

Grid

Site X

Worker nodes

Implementation in AliEn

- Worker nodes keep seeding the packages that they have downloaded
 - Other worker nodes will fetch the content mostly from local nodes
 - Worker nodes from site A are usually firewalled from site B, so no inter-site traffic
 - If initial download is not possible via torrent, fall back to wget and then seed the fetched files
- Multiple versions of the same file can co-exist since they will have different hash codes; old ones will be gracefully phased out.

Current status

- AliEn itself is packaged in a small (35MB) archive
- AliRoot, Root & deps. packaged in single archives: max. 300MB/job
- Subatech is used as testbed
 - LDAP flag to switch modes:
*name=Subatech-CREAM,ou=CE,ou=Services,ou=Subatech,ou=Sites,o=alice,dc=cern,dc=ch
installMethod=Torrent*
 - Production jobs work fine
 - Analysis jobs fail to load a particular library; most probably a configuration issue that is currently tracked
- You can download precompiled packages from <http://alitorrent.cern.ch/>

Future plans

- Full-scale testing of the solution
- Evaluate the need for caching
 - On worker nodes, as files
 - On VoBox, as seeder
 - Regional seeders
 - All these would require managers
- Try to use the solution for distributing data files or pre-compiled PAR files
 - Latest version would be fetched at every execution, no cleanup required for previous ones